DOCUMENT RESUME

ED 182 712 CS 005 268

AUTHOR Jones, Beau F.; Hall, James W.
TITLE Effects of Cross-Classification Strategies for Recalling Prose and for Writing Compare-and-Contrast Essays.
PUB DATE 79
NOTE 28p.; Revised version of a paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 8-12, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Composition (Literary); *Content Reading; Educational Research; *Expository Writing; Grade 7; Junior High Schools; Junior High School Students; *Organization; Reading Research; *Reading Skills; Recall (Psychological); *Retention; *Study Skills

ABSTRACT

After 44 seventh grade students were matched in pairs according to their reading achievement test scores, one member of each pair received training in a cross-classification strategy (matrix outlining), while the other member of each pair received unrelated training. Cross-classifiable information is information that can be categorized along two dimensions simultaneously, such as countries classified in name categories and attribute (type of economy) categories. The trained students had significantly higher scores than untrained students on their free recall of 36 disorganized sentences and on their compare-and-contrast essays. Both tests were given without strategy prompts. The posttest data also indicated that the training was especially effective in increasing learning by low achieving students. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED182712

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Effects of Cross-Classification Strategies for Recalling Prose and for Writing Compare-and-Contrast Essays*

Beau F. Jones

James W. Hall

Northwestern University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Beau F. Jones
James W. Hall
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

*The first author is currently Director of the Chicago Mastery Learning/Learning Strategies Reading (Grades 5-8), Chicago Public Schools, Department of Research and Evaluation, 2021 N. Burling, Chicago, Illinois 60614.

Running head: Matrix Outlining and Analysis

CS005268

EFFECTS OF CROSS-CLASSIFICATION STRATEGIES

FOR RECALLING PROSE AND FOR WRITING COMPARE-AND-CONTRAST ESSAYS

Cross-classifable information is information that can be categorized along two dimensions simultaneously; i.e., into name categories (e.g., names of nations) as one dimension and into attribute categories (e.g., geography, type of government, etc.) as the second dimension. While fiction consists of other types of information, textual information is inherently two-dimensional information (Musgrave & Cohen, 1966) which students from junior high through college must comprehend, memorize, and analyze in oral and written discourse. The present research discusses the development of a complex organizational strategy which seeks to increase comprehension, recall, and analysis of text materials.

In a series of passage organization experiments Frase (1969; 1973) and others (Schultz & DiVesta, 1972; DiVesta, Schultz, & Dangel, 1973; Friedman & Greitzer, 1973; Myers, Pezdek, & Coulson, 1973; Perlmutter & Royer, 1973) studied the effects of different types of passage organization. Using procedures devised by Frase (1969), the researchers in each study constructed a two-dimensional table or matrix in which the name categories were column headings and the attribute categories were row headings. Three types of passages were generated from each matrix: sentences organized by name categories (NO), sentences organized by attribute categories (AO), and randomly organized sentences (RO). In each study the students were given a very brief study period to learn NO, AO, or RO passages. The study period was followed by a free recall test. In addition to type of passage organization, other variables were manipulated: note taking vs no note taking (Schultz & DiVesta, 1972), instructions to organize recall by NO, AO, or free recall for each of the three passages (DiVesta et al, 1973), and instructions to organize recall by NO, AO or free recall given at input or output (Perlmutter & Royer, 1973). It is also important to note that the number of sentences, sentence length, repeated sentence structures, and sematic redundancies between sentences made the name and attribute category relationships very obvious

in the RO condition in some studies (e.g., Myers et al., 1973)

In all of these studies recall and clustering were markedly superior in NO and AO passages. Fortunately, most texts are organized by N or A categories. However, as students study for tests, the input of information from diverse source (lectures, texts, research, labs, etc.) represents RO or something close to it. Yet there is little in the passage organization studies or school curricula that addresses the problem of teaching students to process disorganized information effectively. Jones (1979) reasoned that an effective method of learning disorganized information would be to cross-classify the information to be learned in a two-dimensional table or matrix that would correspond roughly to the matrix from which the sentences were generated. Then, once the information was in a matrix, it would be possible to instruct the students to learn the cell itmes as a series of serial order lists, each list consisting of a name category and the cell items associated with the name, in the order of the attribute categories. Such strategy would thereby utilize both categorizing and serial order strategies.

A similar disparity between task requirements and instruction exists with regard to writing tasks. Students of all ages are constantly asked to analyze N and A information from diverse sources in expository and compare-and-contrast essays (CAC). At the same time, teaching students <u>how</u> to organize and analyze complex, disorganized information is left largely to the teacher's judgement and/or the use of commercially prepared texts which purport to train students in writing skills. Examination of such texts suggests that in fact most, if not all, of them focus on defining and discussing the characteristics of linear outlines and various types of essays, rather than the process of organizing and analyzing information.

Interestingly, one strand of research in writing and thinking paralles the passage organization studies. This research suggests that the most effective means of analyzing complex information to sort it into categories in much the same way that one sorts lists of disorganized words or sentences. And Davis (1976)

in a review of the research on creative thinking and writing, argues that attribute listing and rearrangement (sorting into categories) greatly facilitate the production of various types of analytical and creative statements. Since matrix outlining involves the rearrangment of disorganized attributes into categories, we reasoned that it would facilitate organization and analysis in an essay task.

The task selected to explore this hypothesis was the compare-and-contrast essay. The essence of the compare-and-contrast essay lies in the type of topic sentences or generalizations which are made in the process of comparing (finding similarities among items within the same attribute category) or contrasting (finding differences among the items within the same attribute category). If the categorizing/attribute listing prediction is correct, then students who are trained to outline and analyze information in a matrix should make more creative topic sentences (row generalizations) than students who use traditional linear outlines (clustering strategies) or students who do not outline.

Moreover, the structure of information in a matrix corresponds almost exactly to the paragraph structure in a simple CAC essay. Ideally, a CAC essay should begin with an introductory paragraph that identifies the name and attribute categories to be discussed in the middle section of the essay. This information is in the column and row headings. Paragraph 1 of the middle section should contain a topic sentence that generalizes over all the names to be compared plus details that support the topic sentence. All of the information needed for paragraph 1 is contained in adjacent cells in row 1. This paragraph/row relationship holds for each succeeding paragraph in the essay - not counting the conclusion (which is not required in many schools). Because of this close correpondence between matrix structure and paragraph structure, one would expect that compared to untrained students, trained students would have fewer paragraphs without a topic sentence or row generalization, a classic error, as well as more statement which support the topic sentence.

In the light of these considerations we asked three questions. The first question concerned the memory training. Jones (1979) showed that college students who used cross-classification strategies learned disorganized sentences better than untrained college students. Would it be possible to train younger students to use the same cross-classification strategy to learn disorganized sentences? Second, could younger students be trained to utilize the matrix structure (a) to generate creative/analytical topic sentences for a CAC essay and (b) to organize the overall paragraph structure of a CAC essay? Third, could the students be trained to use such strategies spontaneously; i.e. without being prompted at the time of the test?

Participants and Design

The participants were 44 seventh grade students from a team of three teachers in a suburban junior high school. The students were matched according to a reading achievment test (Gates-MacGinitie) administered by the school to all students one month before the training sessions began. One-half of each matched pair was randomly assigned to the cross-classificatiCn training. The remaining student in each pair was assigned to a second group which received unrelated training (imagery training to learn definitions, Jones and Hall, 1979).

In effect, each group functioned as a control group to the other. The mean reading scores for the trained and untrained groups were 33.11 and 32.11 respectively. It should be noted that the grades which the students received for their records were not the original grades given by the teachers: the actual scores were given only to the experimenter. The grades that each student received were computed on a curve relative to the student's training group.* Thus, this design avoided the moral and motivational problems that arise when only one group is trained for a classroom test. Both groups received the same amount of training and classroom instruction in English, social studies, and science. The measures of interest on the recall test were the total correct and clustering on the recall

* Grading on a curve is the prescribed method of grading in this school.

test. The measures of interest on the writing test were the topic sentence scores, the number of essays with an introductory paragraph, and the scores for paragraph organization. (See below)

Materials and Procedures

Overview. The trained students participated in 10 45-minute training sessions, which were given by the experimenter outside the regular classroom, plus three prompted classroom CAC essay tests given by the teacher. The sessions were distributed over a five month period. This length of time was necessary because of the timing of the various classroom tests. Most training sessions had one focal point: memory training, essay training, or essay transfer training. The memory training involved teaching students to construct and use tables to memorize the information from sets of disorganized sentences. Essay training consisted of instructing the students to write CAC essays using the two-dimensional tables that were constructed as part of the memory training. The essay transfer training consisted of (a) providing some experiences in constructing the name and attribute categories for various hypothetical, subject-related CAC assignments and (b) instructions to construct and use matrices for the three classroom tests.

The 10 training sessions were distributed as follows: session 1 involved categorizing words. Memory training was given in sessions 3, 8, and 9. CAC training using the disorganized sentences was given in sessions 2, 5, and 10. Transfer training was given in sessions 3, 4, 6, and 7. The classroom essay tests were given several days after sessions 2, 5, and 9. The free recall posttest and CAC posttest were administered 6 and 16 days respectively after session 10.

The memory training, the essay training, and the essay transfer training were implemented by a combination of verbal instructions and chalkboard demonstrations (following closely an outline/script prepared by the experimenter) as well as by typed training "handouts." Several features of the training are

important. First, throughout the training sessions the students were constantly actively involved in the learning process by answering fast-paced questions asked by the experimenter: "What is the name category for sentence 1?" "Where does it go on the matrix?" "What is the attribute category for this sentence?" Second, for the recall task, the number of sentences and study time increasingly approximated the target task, namely, to outline and learn 36 disorganized sentences in a 10-minute study period for a free recall test. Third, for each area of training (the memory training, the essay training using disorganized sentences, and the essay transfer training for classroom applications) the strategy prompts were gradually decreased from highly specific, step-by-step prompts, to specific-outline prompts, to general-outline prompts, to no prompts.

Memory and CAC Training Materials. Two sets of categorizable words, five sets of cross-classifiable sentences and three sets of essay-writing exercises constituted the basic content of the training program. The two lists of words consisted of 24 items: six words from each of four categories (e.g., animals). The lists were constructed so that words from the same category were not adjacent. The sentences sets consisted of 9, 20, or 36 sentences. Each set was disorganized such that items from the same name or attribute category were not adjacent. The essay writing exercises utilized the information in the sentence sets.

Classroom Essay Test Materials. In the first classroom training test the students were asked to compare-and-contrast child-care centers in the US and Israel. The test instructions specified that the students had to have at least six "points of comparison or contrast." Two examples (attribute categories) were given as examples. The instructions indicated that the students could use the two categories given as examples. The students were given the two attribute categories to provide an adequate test definition and textual frame or reference for the control students who were beginners at learning to write CAC essays. The step-by-step prompt inserted in the test booklets of the trained students provided a 2 X 6 matrix with the two names and the two attribute categories

in the test instructions supplied in the appropriate places. The prompt further directed the students to complete the remainder of the matrix and then to use it to organize the essay.

The second and third calssroom tests were quite similar. In both tests all students were given the name categories plus all or most of the attribute categories to be discussed. The attribute categories in both of these tests were text specific in the sense that the students were required to discuss the cell items in terms of specific attribute categories which had been given in the text-book or by the teacher in classroom discussions. In classroom test 2 (science) the students were asked to compare-and-contrast three rock types in terms of six attribute categories, three of which were given in the test instructions. In classroom test 3 (English) the students were required to discuss three specific poems in terms of four specified categories (the mother-child relationship, wish fulfillment, the miracle factor, and poetic form). The attribute categories, again, were included to make certain that the control students fully understood the textual frame of reference required for the test.

Training Procedures. During the memory training, the essay training, and the essay transfer training for classroom applications, the strategy prompts were gradually decreased from highly specific step-by-step prompts to specific-outline promtps to general-outline prompts, to no promtps.

1. Step-by-Step Prompts. These involved instructing the students by means of the question-and-answer method with regard to the following: (a) how to construct a matrix so that the number of columns and rows correpsond to the number of name and attribute category headings respectively; (b) how to transfer the information from the sentences to the matrix, working sentence by sentence; (c) how to rehearse the name categories and cell items as lists of related words; (d) how to structure recall so that the items are clustered by names, using ditto marks (Myers et al., 1973; Perlmutter & Royer, 1973) and the attribute items are in serial order. Step-by-step essay prompts were provided in a training handout

which instructed the students how to use a matrix to write a CAC essay. That is, the introductory paragraph contained the CAC question as a statement and identified the name and attribute categories to be discussed in the middle paragraphs. Each middle-section paragraph consisted of a topic sentence (row genernalization) plus supporting details (the cell items) while the concluding paragraph consisted minimally of a summary of the row generalizations. The experimenter also instructed the students to make various types of row generalizations to serve as topic sentences: specification of similarities and differences, the range, or functional equivalence (an underlying common factor in apparently diverse cell items).

2. Specific-Outline Prompts. These consisted of brief oral demonstrations showing the students how to establish the name and attributes category headings for a particular matrix (recall or CAC) plus explicit instructions showing how to fill in the first few cell items.

3. General-Outline Prompts. These consisted of a brief oral review of the steps of a given procedure (recall or CAC) plus the instructions for students to apply the steps to all future test situations for which those procedures were appropriate.

The Essay Posttest. This test was entirely open-ended in that no attribute categories were specified, as is typical in many high school and college assignments. Students were administered one of two posttest questions, counterbalanced according to the matched pairs, such that each pair received the same question but the questions were alternated pair by pair. One question asked the students to compare-and-contrast male and female sex roles. The other asked the students to compare-and-contrast two age roles. Thus, the tests were content-free and differences in achievement would reflect only differences in training not differences in exposure to or attention to given reading materials and class discussion. No attribute categories were specified nor were there any organizational prompts given prior to or at the time of the test. No mention was made of note taking.

The Recall Test. The disorganized sentence booklets were administered as a surprise test by the English teacher and the experimenter jointly to the trained and the untrained students at the same time in the regular classroom. The instructions were reviewed orally by the experimenter. She emphasized that note taking was permitted and that the sentences could be learned and recalled in any order. She also noted that abbreviations, paraphrases, dashes, and incomplete sentences were acceptable as long as the information was correct and complete. The study time was 10 minutes. Recall time was unlimited. There were no prompts before the test or during any part of it.

It should also be noted that a free recall sentence pretest was administered to the untrained students in the early phase of the training period. This was done for two reasons. Pilot research with high school and junior high students indicated that lack of familiarity with test procedures and materials may influence recall. Moreover, it was important to have some measure of the control students' gains in recall without training. Accordingly, the pre- and posttest forms were counterbalanced. The experimental students received the same two tests (counterbalanced) in session 9 and for the posttest. Session 9 and the posttest were five weeks apart.

Scoring and Grading. The free recall test and all training exercises for the free recall test were scored for total correct and clustering by the expeimenter. The formula used to measure clustering was that used by Frase (1969): $R/(T-K) \times 100$, where R = the number of repetitions from the same category; T = total number of correct sentences; and K = the number of categories in recall. The three CAC classroom training tests were graded for overall quality by their respective teachers. These scores were checked by the experimenter and external reader who did not know which students were in each group. The reader then scored each paper separately for the presence of a topic sentence for each paragraph, the quality of the topic sentence (using a scale of 1-3 with clearly specified criteria), the amount of information (attributes) presented in support of each topic sentence,

and overall quality.

Results: Posttests

The mean scores for the sentence recall test and the CAC essay are given in Table 1. One-tailed t-tests were conducted on all of the measures in Table 1 except for the presence of an introductory paragraph. (See below.) The difference between the means on the recall test for total correct and clustering were significant: $t(21) = 2.83$, $p < .005$ and $t(21) = 2.23$, $p < .05$. The large standard deviations for total correct for the trained and untrained students (SD's = 6.82 and 5.32 respectively) suggest that there were substantial variations in performance. Examination of the raw scores indicated that there were two different patterns of recall for the two groups. The scores for the trained students were bimodal: The mean score for those above the median was 23.30 whereas the mean score for those below the median was 15.18. In contrast, the mean recall scores above and below the median for the untrained students were 15.72 and 12.50 respectively.

Insert Table 1 chart here.

The difference between the means for the number of paragraphs with a topic sentence, the quality of the topic sentence, and the information given to support the topic sentence were also significant: $t(19) = 2.79$, $p < .01$; $t(19) = 2.16$, $p < .05$; and $t(19) = 2.63$, $p < .01$ respectively. The mean scores for the trained and untrained students above and below the median were not bimodal on any of the CAC measures; (e.g., for number of paragraphs with a topic sentence, the mean score above the median was 1.9 compared to 2.1 for those below the median.) Curiously, only three of the trained students and one of the untrained students had introductory paragraphs or statements that were separated from the body of the essay. Consequently, no statistical analyses were performed on this measure, and no students had concluding paragraphs.

While statistical comparisons of high and low achievers were not appropriate, due to regression effects, it should be noted that the low achievers in the trained group (those below the median in reading) were consistantly equal or markedly superior to the high achievers in the control group (those above the median in reading) for each of the measures analyzed the Table 1. The mean recall scores for the trained low achievers was 15.88, compared to 15.72 for the untrained high achievers. The mean clustering score for the trained low achievers was .52, compared to .51 for the untrained high achievers. The mean scores for number of paragraphs with topic sentences, quality of topic sentences, and information were 2.10, 4.40 and 14.10 respectively for the trained low achievers, compared to 1.30, 2.80, and 12.90 for the untrained high achievers.

Because the note-taking procedures used by the control students on the recall test were highly variable, no statistical analyses were performed on this data. However, there were 16 control students who took notes, and all but three of them made matrixes. In contrast, only eight of the trained students made matrices for the CAC essay. Postlearning interviews indicated that many students did not bother to make the matrices on their test because, with only two name categories, an appropriate matrix could be visualized in ones' head without writing it down on paper.

Interestingly, the control student made significant gains in clustering between the pretest and posttest. The pretest mean was .27 (SD = .19), compared to the posttest mean .46 (SD = .23); $t(21) = 3.82$, $t < .005$. With regard to the control students' gains between pre- and posttest on the recall test, the mean for the total correct for the pretest was 10.02 (SD = 4.65) compared to 13.23 (SD = 5. 32) on the posttest. This difference was not significant. Thus, while the control students apparently learned to categorize their recall without training, they were not able to increase recall without training.

Results: The CAC Essay Test (Overall Grades)

Perhaps the most interesting results of all were the (holistic) classroom

grades of the classroom training tests. Recall that the first test involved a step-by-step prompt inserted inside the test booklets of the trained students. The range of unadjusted grades given by the teachers and external reader to the students ranged from A- to A. The grades of the untrained students' tests consisted largely of B's and C's. While we had expected substantial differences between the trained and untrained students, we had not expected the uniformity of excellent grades among the trained students. Neither apparently, had the team teachers, who were amazed that any method of teaching could have such an immediate and marked effect on students required to do a novel and complex task. Moreover, the fact that the immediate and marked effect was present for the low achieving students as well was cause for considerable discussion among the teachers.

Since the trained students had only specific-outline prompts and general-outline prompts on the second and third classroom training tests, it was expected that there would be differences in the classroom grades between the trained and untrained students but that these differences would be less than those obtained when the step-by-step prompt was given to the trained group. Surprisingly, this was not the case. The science essay test grades for both groups were uniformly high; M= 3.7 and 3.6 for the trained and untrained students respectively. Similarly, the mean grades for the trained and untrained students for the English test were 3.9 and 3.7 respectively. Both the science and English teachers thought that these average grades scores were exceptionally high, compared to those achieved in the past. This judgement was confirmed by the principal and two other teachers who read the essays, as well as by the external reader.

In contrast, the classroom grades given by the teachers and external reader for the posttest essays were unexpectedly low for all students. More specifically, while there were significant differences between the two groups in the quality of the topic sentences, the number of paragraphs with topic sentences, and amount of supporting information, all of the analytical scores were skewed toward the lower end of the grading scales rather than the upper end, as in the previous tests. Moreover, the total number of paragraphs was far fewer than on any previous essay.

These differences between tests were noted by the external reader and all three team teachers. In effect, the cross-classification strategy training facilitated organization and the production of specific types of analytical statements but not overall quality when the questions contained no attribute categories and there were no prompts.

Perhaps the most interesting findings that arose from this study were these unanticipated results concerning the CAC essay tests. The pattern of results of the three training tests and the posttest may be explained as an interaction of sets of related factors: (a) the degree of explicitness of the attribute categories in the test and (b) the degree of explicitness of the prompts. This interaction is summarized in Fig. 1 below. The highest classroom grades were attained when the attribute categories were explicit in the test questions or when the prompts were explicit. Obversely, the lowest grades were obtained by the trained and untrained students when the attribute categories and the analytical procedures were not specified in the prompt or test question.

Insert Figure 1 about here.

Results: Survey Data and Postlearning Comments

After the recall posttest all students were administered a questionaire to assess how the students felt about the training program. The four students who were the lowest in reading comprehension stated that they found the procedures difficult to learn and therefore not beneficial. Apart from these comments, the comments of the remainder of the trained students were remarkably positive. All of them felt that the CAC procedures were either easy or moderately easy to learn and that the training was most beneficial in helping them to organize CAC essays. The recall training was difficult but worth knowing, though several students commented that the recall procedures might best be taught in high school since seventh graders are seldom required to memorize cross-classifiable information.

There are many anecdotes that we could relate regarding the perceived impact of the training on the students, but the ones that were most interesting were the comments made by various students during the year after the study. Particularly amusing was the comment of one of the students who called up the experimenter a year later to say that while she had thought that the recall training was totally unnecessary the preceding year, she wanted to make certain that she remembered the procedures so that she could use them to study for biology tests. Another youth stated flatly, much to the chagrin of his teacher, that this training was "the most important thing that he had learned last year." But the most surprising comments came from the number of parents who told us how much their children had benefitted from the training during and after the experiment. And, several of the students, it seems, had even taught their parents the procedures!

Discussion

Undoubtely, the length of the study, the complexity of the training program, and the few instances of lack of control of the strategy training information limit the inferences that one can made from this study. In spite of these limitations, the data are suggestive. First, the marked differences between the trained and untrained students on the free recall sentence posttest indicate that it is possible to train seventh graders to use the same clustering and serial order strategies to facilitate recall that were taught to the college students in the earlier Jones (1979) study. The CAC essay posttest data indicate that seventh graders can also be trained to use cross-classification strategies to increase the production of analytical statements as topic sentences, the production of paragraphs with topic sentences, and the production of information to support the topic sentences but not the production of introductory and concluding paragraphs.

The posttest differences between the trained and untrained students are consistant with any number of developmental studies (e.g., Moely & Jeffery, 1974) which have found marked increases in attainment when younger students are taught to use the strategies employed by older students. What is interesting in the

present research in that the organizational strategies learned by the younger students were highly complex strategies applied to two very difficult tasks.

The gains in clustering among the untrained students between pre- and posttest are also consistent with these developmental studies. Whatever the source of these gains (exposure to relevant classroom experiences, maturation or exposure to strategy training information), increased strategy use without prompting over a 5-month period would be predicted from developmental studies and also, not suprisingly, from an examination of the junior high objectives in this school and other good schools; i.e., many good junior high schools focus on teaching their students to organize what they write in terms of name or attribute categories.

Unfortunately, no clearcut statement can be made with regard to the question of whether or not the students use the strategies spontaneously. Although the posttest was given four days after the last training session without any prompts, the presence of the experimenter and the free recall testing procedures (which are never used in schools) may have functioned as cues to use the strategies. Perhaps the best that can be said on this point is the most of the trained students used the strategies effectively without explicit prompting. This problem did not exist with regard to the writing posttest, which is a traditional type of test in junior high and which was administered by the teacher alone. Consequently, it may be assumed that the strategic behaviors employed by the students in this situation were entirely unprompted.

While this relationship between the degree of explicitness of the test question or prompt and grades was not predicted, in retrospect, such relationship hardly seems surprising. Apparently, the provision of attribute categories in the test question functioned as a cue for the production of analytical statements (topic sentences).

There is a hypothesis that explains the pattern of results of the CAC essay tests and the free recall test: the encoding variability hypothesis.

This hypothesis was first advanced by Melton (1967) and Martin (1986). They argue that encoding information through different contexts increases the amount of information associated with an item that could serve as retrieval cues, and therefore the probability that the item may be recalled. The evidence for this hypothesis, according to Shimmerlik (1978), was based on research comparing the benefits of space vs massed repetitions in memory. Melton and Martin reasoned that space repetitions are recalled better than massed repetitions because the former have a greater variability of context. In a juxtaposition of the Melton/ Martin research and the passage organization studies, Shimmerlik used the encoding variability hypothesis to explain some unexpected findings of DiVesta, Schultz, & Dangel (1973). They predicted that recall would be superior when the passage organization and the required recall organization (N or A) were congruent. When recall was highest in the incongruent condition, they accounted for this result in terms of the added information concerning passage organization in that condition. Shimmerlik argued that the DiVesta et al findings could be the result of increased contextual associations (encoding variability) that were attained in the process of reorganization.

Shimmerlik's argement may explain why matrix outlining facilitates recall. Speaking broadly, it may be argued that encoding information through two different organizations (N an A) provides greater contextual associations than encoding through one organizational dimension, and therefore would increase recall. More specifically, outlining the information in a matrix which systematically utilizes N and A categories, simultaneously associates each cell item with 2-4 related items. Each inner cell item is adjacent to (and therefore associated with) two items within the same name category (the items above and below the cell) and two items within the same attribute category (the items on either side). The instructions to memorize the cell items as a series of serial order lists (each consisting of the column heading plus the cell items in each column) gives two additional associations for each item: the column heading and the row heading.

Thus, the outlining procedures and the instructions, if followed exactly, generate 2-4 cell associations plus 2 heading associations for each item. These associations are far greater in number than those in a list of sentences or in a linear outline.

Shimmerlik's argument may be extended to explain the expected and unexpected essay test results. First, the expected results -- the uniformly high grades of the trained students when they were administered a step-by-step prompt on the first transfer training test. Matrix outlining and analysis procedures provide not only a literal but an inferred contextual vairability. That is, these procedures force the students (a) to relate each attribute item to a specific N and A category (the literal contextual variability), and (b) to make an inference about the similarity or differences among the attribute items within a given attribute category (the inferred variability) which links all of the row items by definition. Moreover, we assume that physically embedding the cell items in adjacent columns and rows facilitates the inference process for several reasons: writing down the items to be analyzed ensures that they will not be forgotten and allows the students to give their entire attention to the inference process. having the items physically adjacent (rather than having no notes or spreading them out over one or more pages in a linear outline) ensures that the inference process is not interrupted by the need to find an item at a different location, and thereby increases the time available to scan adjacent cells. It is possible that the physical proximity of the written items may generate contextual associations that would not have otherwise existed.

Paradoxically, this same line of reasoning may also explain the unexpected CAC training essay results - the high grades for the untrained students when the attribute categories were specified in the question for the second and third training tests. In each of these test situations there were only three name categories to be compared within each attribute category. The English test, for example, required analysis of only three poems (the name categories) within four attribute categories. Given the small number of N categories, specifying the N

and A categories in the test question would have virtually the same effect as matrix outlining. With only the three poems, the students needed to recall and analyze only three sets of information for each paragraph; e.g., the mother-child relationship in each of the three poems. This could be done without notes of any kind. Moreover, it is likely that specifying the attribute categories in the test question generates in one's mind numerous contextual associations without writing out the facts - if the N categories are few in number. A larger number of N categories would, of course, require a greater memory load - not only of the attribute items but also the contextual associations for each one. Thus, it seems likely that explicit CAC essay questions - i.e., those which are limited in the number of categories (three or less) and which specify the attribute categories -- may generate the same level of contextual associations to aid the production of analytical statements that matrix outlining does. If this is true, providing explicit CAC essay questions may influence grades quite independently of matrix outlining instructions.

Finally, the encoding variability hypothesis may account for the unexpected, overall low-quality essay posttest grades. The posttest, it will be recalled, had no attribute categories in the test question. The effects of this were twofold. First, the students were forced to generate their own attribute categories. With no guidelines as to the number and content, the attribute categories in the essays in both groups were very obvious and not as conducive to stimulating contextual associations as those provided by the teacher had been. Second, the attribute categories were limited in number with only 2-3 per essay, compared to 4-6 in previous essays. Fewer attribute categories mean fewer contextual associations, fewer analytical statements, and fewer paragraphs.

Evidence to support this line of reasoning comes from the comments of the trained students who did not construct matrices for the second and third training essays. In effect, they argued that constructing a matrix on paper was not necessary because the N categories were few in number and A categories were specified. As one student said, "I can work out the topic sentence in my head, write it down,

and write my examples down, paragraph by paragraph." Unfortunately, the untrained students were never interviewed on this point. However, the greatest strength of the encoding variablilty hypothesis, does not arise from direct evidence but rather from the fact that it neatly accounts for the recall data and the essay data both expected and unexpected.

To summarize, significant differences were found in the recall and CAC essay posttest results as expected. Two sets of findings were unexpected: the uniformly high grades of the untrained students on the second and third essay training tests and the poor quality of overall grades for both the trained and the untrained students for the essay posttest. Both the expected and the unexpected results were found to be a function of the degree of explicitness in the prompts or in the specification of attribute categories in the test question. Both the expected and unexpected results were explained in terms of the encoding variability hypothesis.

Implications for Future Research

Many aspects of the present study suggest that further research is warranted. Combining matrix outlining with instructions to use clustering and serial order strategies for recall appears to be a most fruitful approach to help younger students memorize extensive amounts of cross-classifiable textual prose. Moreover, we seem to have discovered not one, but two methods of obtaining very high achievement in writing: the explicit step-by-step prompt, and the explict CAC essay test question. Also important, most of the students saw quickly the relevence and usefulness of the CAC training, and viewed participating in the experiment as a beneficial and enjoyable experience.

Although the results were encouraging for practical reasons, the analysis of both the expected and unexpected results indicate that there are several methodological questions that will have to be considered in subsequent studies - both basic and applied. Especially important are: What constitutes a fair control group? To what extent is it possible to separate out the effects of different parts of

the training procedures? What is an appropriate CAC posttest? Should it be content free or not? What are the effects of specifying the attribute categories in the CAC essay training test? Is it possiple to develop a recall task that does not cue the learner to use the prescribed strategy? And what about scoring? Although the scoring procedures used in the present study were reasonably objective, the method was time-consuming and did not cover the full range of items that classroom teachers typically use in grading. Would it be possible to develop an anlytical scoring system that was not so time-consuming and yet covered the full range of items that teachers grade for?

Another set of issues relates to maximizing control over what is taught and how classroom tests are administered. Training students to use matrix outlining with recall strategies and/or writing strategies is, of necessity, a long-term process. Developing instructional "scripts" to go with practice exercises and extensive answer keys that the experimenters or teachers themselves could use would greatly aid control over instructional information as well as feedback on correction of errors. But even with the addition of highly structured materials, monitoring teacher use of the instructional materials and administration of classroom tests would also be important.

The theoretical issues that could be investigated are diverse. First, the encoding variability hypothesis suggests that spatial arrangement of the information to be learned or analyzed may be a critical variable in recall and in writing CAC essays. If this is true, it would be interesting to compare sentence recall using matrix outlining vs linear outlining which separates items that would be spatially linked in a matrix. Conversely, one could test the encoding variability hypothesis itself by comparing information in a linear outline that is spread out over many pages to information in a linear outline that has all of the information on adjacent lines.

Second, several lines of research on strategy use indicate that at about sixth grade, high achieving students spontaneously begin to use a range of effective learning and reading strategies with increasing efficiency through high school

years; low achieving students do not make this developmental shift at any time without intervention (See Rohwer, 1971 and Jones, Monsaas, & Katims, 1979 for reviews.) The students in the present study were high achievers in terms of their reading and vocabulary scores on standardized measures. The gains in clustering on the recall test made by the control students suggest that there may be a developmental shift in the use of clustering and serial order strategies. If high and low achievers from grades 5, 7, 9, and 11 were given a disorganized prose passage task, would there be individual differences with regard to age and achievement level such as that indicated by the strategy research? Would the same pattern of individual differences emerge from a developmental study of the CAC essay task? Would low achieving students benefit from strategy training on both tasks?

Third, two sources of data suggest that there is considerable overlap in the cognitive skills involved in reading, recalling, and analyzing complex prose. The high correlation between the reading comprehension scores, clustering, total recalled, and essay scores (unpublished data) and suggest similar skills. Specifying the attribute categories in the essay test question functioned as a cue in the same way that providing name or attribute categories functions as a cue in a recall test. If there is such an overlap, it is likely that other variables which are known to influence the recall of cross-classifiable information may also influence analysis of such information; e.g., the existence of semantic redundancies in the test (Frase, 1973).

Research questions on the effects of matrix outlining and analysis for essays largely parallels questions regarding the recall task. Clearly, the first research problem is to trace out the effects of the component parts of the training and test situation; namely, the effect of specifying the A categories in the test question and the effects of different numbers of N categories with and without matrix outlining. The data from such studies would not only clarify critical methodological and pedogogical considerations but would also relate to the encoding

variability hypothesis. It would also be useful to consider individual difference vaiables and the effects of matrix vs linear outlining procedures for reasons given above.

Another line of research involves establishing the conditions which best facilitate application to classroom tasks. One approach would be to develop new materials to solve three practical problems which may have limited the extent of application in the present study: (a) the limitations of the reader's world knowledge of appropriate attribute categories and criteria for selecting them when they are not given in the test question, (b) the limitations of the outline structure and procedures which allowed only for putting attribute details or facts in each cell (essentially only a two-level outline consisting of category headings and details), (c) the limitation of the outline structure and procedures which provide only for the production of several types of analytical statements to use a topic sentence. Currently, we are developing materials which address these limitations as well as the problem of control over instructional information and the scoring problems mentioned earlier.

References

Davis, G. Research and Development in Training Creative Thinking. In J. R. Levin & V. L. Allen (Eds.), Cognitive Learning in Children. New york: Academic Press, 1976.

DiVesta, F. J., Schultz, C. B., & Dangel, T. R. Passage organization and imposed learning strategies in comprehension and recall of connected discourse. Memory and Cognition, 1973, 1, 471-476.

Frase, L. T. Paragraph organization of written materials: The influence of conceptual clustering upon the level and organization of recall. Journal of Educational Psychology, 1969, 60, 394-401.

Frase, L. T. Integration of written materials. Journal of Educational Psychology, 1973, 65, 252-261.

Friedman, M., & Greitzer, F. Organization and study time in learning from reading. Journal of Educational Psychology, 1972, 63, 609-616.

Jones, B. F. Individual differences in the use of diverse learning strategies and achievement in high school. (Doctoral Dissertation, Northwestern University, 1976). Dissertation Abstracts International, 1976, 37, 7040A. No. 77-10, 042.)

Jones, B. F. Effects of teaching cross-classification strategies for recall of disorganized prose passages. Paper submitted to the Journal of Applied Developmental Psychology for consideration for publication, April, 1979.

Jones, B. F., Monsaas, J. A., Katims, M. Improving reading comprehension: Embedding diverse learning strategies within a mastery learning instructional framework. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1979.

Martin, E. Stimulus meaningfulness and paired - associate transfer: An encoding variability hypothesis. Psychological Review, 1968, 75, 421-441.

Melton, A. W. The situation with respect to the spacing of repetitions and memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 596-606.

Moeley, B.E,, & Jeffery, W. E. The effect of organization training on children's free recall of category items. Child Development, 1974, 45, 135-143.

Musgrave, B. S., & Cohen, J. C. Relationships between prose and list learning. In E. Z. Rothkopf & P. E. Johnson (Eds.) Verbal Learning and the Technology of Written Instruction. New York: Teacher's College Press, 1971.

Myers, J. L., Pezdek, K., & Coulson, D. Effect of prose organization upon free recall. Journal of Educational Psychology, 1973, 65, 313-320.

Perlmutter, J., & Royer, J. M. Organization of prose materials: Stimulus, storage and retrieval. Canadian Journal of Psychology, 1973, 27 200-209.

Rohwer, W. D. Jr. Elaboration and learning in childhood and adolescence. In H. W. Reese (Ed.), Advances in Child Development (Vol. 8). New York: Academic Press, 1973.

Schultz, C. B. & Di Vesta, F. J. Effects of passage organization and note taking on the selection of clustering strategies and on recall of textual material. Journal of Educational Psychology, 1972, 63, 244-252.

Shimmerlik, S. M. Organization theory and memory for prose: A review of the literature. Review of Educational Research, 1978, 1, 103-120.

		Explicitness of Attribute Categories	
		High	Low
Explicitness of the Training Prompts	High	Not Tested	Social Studies Training Test #1 Trained = A Untrained=Bs/C's
	Low	Science & English Training Tests 2 + 3 Trained = A untrained = A	Content-Free Posttest Trained = B-/C+ untrained = C-/D+

Frg. 1. Compare-and-contrast essay average grades as a function of the degree of explicitness of the attribute categories in the test question and the degree of explicitness of the prompt.

Table 1

Means and Standard Deviations for the Postest Subtests

Measure	Control		Trained		t
	M	SD	M	SD	
Recall of Sentences (22)[a]					
Number Correct	13.23	5.32	18.75	6.82	2.83***
Clustering[b]	.46	.23	.60	.22	2.23*
Final CAC Essay (20)[a]					
Introductory Paragraph[c]	.05	___	.15	___	___
Paragraphs with Topic Sentences	1.25	.72	2.05	1.23	2.79**
Quality of Topic Sentences	2.50	1.50	4.10	2.71	2.16*
Information Supporting Topic Sentences	11.55	3.25	14.70	4.39	2.63**

[a]Numbers in parentheses represent the number of pairs of students who were tested.

[b]Cluster scores represent the proportion clustered for name and attribute clustering combined.

[c]See text.

* $p < .05$

** $p < .01$

*** $p < .005$